

# UCLA

## UCLA Previously Published Works

### Title

International network of cancer genome projects.

### Permalink

<https://escholarship.org/uc/item/4409c5hv>

### Journal

Nature, 464(7291)

### ISSN

0028-0836

### Authors

International Cancer Genome Consortium  
Hudson, Thomas J  
Anderson, Warwick  
et al.

### Publication Date

2010-04-01

### DOI

10.1038/nature08987

Peer reviewed



Published in final edited form as:

Nature. 2010 April 15; 464(7291): 993–998. doi:10.1038/nature08987.

## International network of cancer genome projects

### The International Cancer Genome Consortium\*

#### Abstract

The International Cancer Genome Consortium (ICGC) was launched to coordinate large-scale cancer genome studies in tumors from 50 different cancer types and/or subtypes that are of clinical and societal importance across the globe. Systematic studies of over 25,000 cancer genomes at the genomic, epigenomic, and transcriptomic levels will reveal the repertoire of oncogenic mutations, uncover traces of the mutagenic influences, define clinically-relevant subtypes for prognosis and therapeutic management, and enable the development of new cancer therapies.

The genomes of all cancers accumulate somatic mutations<sup>1</sup>. These include nucleotide substitutions, small insertions and deletions, chromosomal rearrangements and copy number changes that can affect protein-coding or regulatory components of genes. In addition, cancer genomes usually acquire somatic epigenetic “marks” compared to non-neoplastic tissues from the same organ, notably changes in the methylation status of cytosines at CpG dinucleotides.

A subset of the somatic mutations in cancer cells confers oncogenic properties such as growth advantage, tissue invasion and metastasis, angiogenesis, and evasion of apoptosis<sup>2</sup>. These are termed “driver” mutations. The identification of driver mutations will provide insights into cancer biology and highlight novel drug targets and diagnostic tests. Knowledge of cancer mutations has already led to the development of specific therapies, such as trastuzumab for HER2/neu positive breast cancers<sup>3</sup> and imatinib, which targets BCR-ABL tyrosine kinase for the treatment of chronic myeloid leukemia<sup>4,5</sup>. The remaining somatic mutations in cancer genomes that do not contribute to cancer development are called “passengers”. These mutations provide insights into the DNA damage and repair processes that have been operative during cancer development, including exogenous environmental exposures<sup>6,7</sup>. In most cancer genomes, it is anticipated that passenger mutations, as well as germline variants not yet catalogued in polymorphism databases, will substantially outnumber drivers.

Large-scale analyses of genes in tumors have revealed that the mutation load in cancer is abundant and heterogeneous<sup>8–13</sup>. Preliminary surveys of cancer genomes have already demonstrated their relevance in identifying new cancer genes that constitute potential therapeutic targets for several types of cancer, including PIK3CA<sup>14</sup>, BRAF<sup>15</sup>, NF1<sup>10</sup>,

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Author Information: Correspondence should be addressed to T.J.H. (tom.hudson@oicr.on.ca).

\*List of participants and affiliations appear at the end of the paper.

KDR10, PIK3R19, and histone methyltransferases and demethylases<sup>16,17</sup>. These projects have also yielded correlations between cancer mutations and prognosis, such as IDH1 and IDH2 mutations in several types of gliomas<sup>13,18</sup>. Advances in massively parallel sequencing technology have enabled sequencing of entire cancer genomes<sup>19-22</sup>.

Following the launch of comprehensive cancer genome projects in the United Kingdom (Cancer Genome Project)<sup>23</sup> and the United States (The Cancer Genome Atlas)<sup>24</sup>, cancer genome scientists and funding agencies met in Toronto (Canada) in October 2007 to discuss the opportunity to launch an international consortium. Key reasons for its formation were: (1) the scope is huge; (2) independent cancer genome initiatives could lead to duplication of effort or incomplete studies; (3) lack of standardization across studies could diminish the opportunities to merge and compare datasets; (4) the spectrum of many cancers is known to vary across the world; (5) an international consortium will accelerate the dissemination of datasets and analytical methods into the user community.

Working groups were created to develop strategies and policies that would form the basis for participation in the ICGC. The goals of the Consortium (Box 1) were released in April 2008 ([http://www.icgc.org/files/ICGC\\_April\\_29\\_2008.pdf](http://www.icgc.org/files/ICGC_April_29_2008.pdf)). Since then, working groups and initial member projects have further refined the policies and plans for international collaboration.

## Bioethical Framework

ICGC members agreed to a core set of bioethical elements for consent as a precondition of membership (Box 2). The Ethics and Policy Committee has created patient consent templates for both prospective collection and retrospective use of samples and data for ICGC projects. Differences in project-specific requirements and national legal frameworks may require some local amendments, while still reflecting the core principles of ICGC.

The ICGC recognizes a delicate balance between protecting participants' personal data and sharing these data to accelerate cancer research. Data access policies have been drawn up that are respectful of the rights of the donors, while allowing ICGC data derived from samples to be shared ethically among a wide research community. Two levels of access have been implemented. For data that cannot be used to identify individuals, "Open access" datasets are publically available. These include data such as gender, age range, histology, normalized gene expression values, epigenetic datasets, somatic mutations, summaries of germline data, and study protocols. "Controlled access" datasets contain germline genomic data and detailed clinical information that are associated to a unique individual whose personal identifiers have been removed. To access controlled datasets researchers must seek authorizations by contacting the Data Access Compliance Office (DACO) (<http://www.icgc.org/daco>). An independent International Data Access Committee (IDAC) oversees the work of the DACO and provides assistance with resolving issues that arise.

## Pathology and Clinical Annotation

Large-scale genomic studies of human tumors rely on the availability of fresh frozen tumor tissue. To address the paucity of samples that meet ICGC standards, many projects have

initiated prospective collections of high quality source material. Accordingly, the ICGC recommended procedures to promote consistency of sample processing throughout the Consortium and ensure a series of quality features such as high tissue integrity and tumor cell content. Each project will need to include diverse data types such as environmental exposures, clinical history of participants, tumor histopathology, and clinical outcomes.

Tumors display considerable clinical and biological heterogeneity which has resulted in a variety of tumor classifications. Within the ICGC, special measures are taken to promote the consistency of diagnosis. These include the coordination of diagnostic criteria among groups investigating tumors that are related, and policies that all samples will be reviewed by at least two independent reference pathologists. Furthermore, images of the stained tumor sections (or blood smear or cytospins for hematological neoplasias), from which diagnoses were made, will be stored and made available to the community.

Although different tumor types may require specific procedures for tumor acquisition or compilation of clinical and environmental data, ICGC has set guidelines regarding the use of common definitions and data standards. This will allow ICGC data users to identify correlations between tumor-specific molecular changes with clinical and histopathological data including prognosis, prediction of therapy response and tumor classification schemes for diagnosis.

## Study Design and Statistical Issues

To identify cancer-related genes, one needs to detect genes that are mutated at a higher frequency than the background mutation rate. Given that several driver genes have been found to be mutated at low frequencies, ICGC will identify somatic mutation observed in at least 3% of tumors of a given subtype. ICGC determined that 500 samples would be needed per tumor type (although for rare tumor types, a smaller sample size may be justified). In practice, the degree of heterogeneity of a given tumor type is difficult to know in advance, such that some particularly heterogeneous tumor types may require larger sample collections.

## Cancer Genome Analyses

High-quality catalogues of somatic mutations from whole cancer genomes will ultimately be the ICGC standard. Shotgun sequencing employing second generation technologies can detect all classes of somatic mutation implicated in cancer. Moreover, if the level of coverage is sufficient, comprehensive high quality catalogues of somatic mutations from individual cancer genomes can be acquired with >90% sensitivity and >95% specificity. In order to achieve this, it will be necessary to sequence both the genome of the cancer and of a normal tissue from the same individual to distinguish germline variants. Although a few genomes of this standard have already been generated, the cost and the continuing technology development will mean that interim analyses of particularly informative sectors of the genome will be carried out, for example of all coding exons and microRNAs.

For each individual cancer genome, the catalogue of somatic mutations will be supplemented by genome-wide information on the state of methylation of CpG

dinucleotides. The optimal strategies and technologies to achieve this are not yet clear. Moreover, the genomes of individual cancers will be accompanied, where possible, by analyses of the transcriptome. Although conventional array-based approaches currently predominate, it is preferable that RNA sequencing becomes the standard as sequencing has a greater dynamic range<sup>25</sup> and provides additional information including novel transcripts and sequence variants<sup>26</sup>.

## ICGC Datasets

The distributed nature of the Consortium coupled with the large size of the datasets makes it cumbersome to store all data in a single centralized repository. For this reason, the ICGC has adopted a “franchise” database model for integrating the information and making it available to the public. Under this model, each member project releases tumor information by copying it into its local franchise database after it has been quality checked. Each franchise database shares a common schema to describe the specimens, the associated clinical information, and their genome characterization data. ICGC primary data files, including sequencing traces, are sent to the National Center for Biotechnology Information (NCBI) and/or the European Bioinformatics Institute (EBI) for archiving, while interpreted data sets, such as somatic mutation calls, are stored in franchise databases. The ICGC franchise databases and web portal use BioMart<sup>27</sup>, a data federation technology originally developed for use in Ensembl<sup>28</sup>, and since adopted for use by multiple model organism and genome databases. The management of the ICGC data flow is the responsibility of the ICGC Data Coordination Center (DCC) located at the Ontario Institute for Cancer Research.

The DCC also operates the ICGC data portal which allows researchers to access both Open and Controlled access portions of the ICGC data. The portal provides a variety of user interfaces that range from simple gene-oriented queries (“Show me all the non-silent coding mutations identified in PIK3R1 for all cancers.”) to queries that integrate genomic, clinical, and functional information (“Show me all members of the toll receptor pathway having deletions in stage III breast cancer.”). These queries will be distributed across the franchise databases in a manner that is invisible to the user. The portal will also provide links to the primary files at NCBI and EBI, interfaces for generating tabular reports, data dumps in common bioinformatics formats, and other visualizations including genome browser tracks, pathway diagrams and survival curves. The portal is available via a link at <http://www.icgc.org>.

At the time of this publication, the following cancer and reference datasets will be available through the ICGC web portal:

- Initial data releases from ICGC members for breast cancer (UK), liver cancer (Japan), and pancreatic cancer (Australia and Canada);
- A whole genome dataset of a metastatic melanoma cell line (COLO829)<sup>6</sup>;
- Open datasets from the TCGA for glioblastoma multiforme (GBM) and serous cystadenocarcinoma of the ovary (see below);

- Whole exome somatic mutation data from 68 individuals with breast, colorectal, pancreatic cancer and GBM11-13;
- Links to the human reference genome (<http://www.genomereference.org/>) and gene annotations from the GENCODE Project (<http://www.sanger.ac.uk/gencode/>) which includes the CCDS gene set<sup>29</sup>;
- Links to dbSNP<sup>30</sup> and the HapMap<sup>31</sup> databases, providing access to common patterns of variation in reference population samples;
- Links to Reactome<sup>32</sup>, a curated database of biological pathways in human;
- A set of reference gene models, mirrored from ENSEMBL<sup>28</sup>.

The current version of the web portal provides an entry point to the open access data tier via interactive query as well as bulk download of data files. We expect that in mid 2010 both open access and controlled data will be available.

The ICGC recently established a bioinformatics analysis working group to compare pipelines, analytic methods, consistency within and among algorithms, and establish guidelines or best practices for the Consortium. Over time, significant resources will be deployed to develop strategies to analyze the large complex datasets generated by ICGC member projects, and provide value-added views of cancer genomic data by integrating them with other biological and epidemiological datasets.

## Data Release and IP Policies

The data release policies of the ICGC are intended to maximize public benefit while, at the same time, protecting the interests and rights of sample donors and their relatives. Members of the ICGC are committed to the principles of rapid data release (with appropriate controlled access mechanisms), in concordance with the Toronto Statement<sup>33</sup>. ICGC members encourage the scientific community to use any data that targets specific genes and mutations, without any restrictions. In order to allow ICGC members the opportunity to be the first to publish global analyses from datasets they generate, the Consortium has also agreed that member projects may specify conditions that include a time limit during which other data users are asked to refrain from publishing global analyses (defined by several ICGC member projects as 100 tumors and matched controls), a provision referred to as a “publication moratorium”. In order to allow time for a dataset to be analyzed and submitted for publication, ICGC members will have at most one year after released datasets reach the specified threshold before third parties are permitted to submit manuscripts describing global analyses. Further details on data release guidelines for data producers, users and reviewers are available <http://www.icgc.org>. Users of ICGC data are expected to respect these terms and to cite this manuscript and the source of pre-publication data, including the version of the dataset. In cases of uncertainty, scientists using ICGC data are encouraged to contact the member projects to discuss publication plans.

ICGC members believe that maximum public benefit will be achieved if the data remain publicly accessible without patent restrictions hence no claims to possible intellectual property (IP) derived from primary data (including somatic mutations) will be made. Users

of ICGC data (including ICGC members) may elect to perform further research and to exercise their IP rights on these downstream discoveries. If this occurs, users are expected to implement licensing policies that do not obstruct further research.

## Initial ICGC Projects

Currently nine countries and two European consortia have initiated cancer genome projects under the umbrella of the ICGC. The initial projects, listed in an online table that accompanies this article, will analyze tumor types found around the globe and throughout the human body affecting a diversity of organs including blood, brain, breast, kidney, liver, pancreas, stomach, oral cavity, and ovary. Over time, the ICGC will investigate fifty or more types and subtypes of cancer in adults and children. In the case of tumors with multiple subtypes, analyses should be focused on subtypes that may be defined on pathological, molecular, etiological or geographical differences. It is expected that some cancer types will be studied in parallel in different parts of the world, as the mutation profiles may differ among populations. The consortium has enabled the coordination of initial projects analyzing similar cancers in different countries, and in some cases, the redirection of resources to launch new projects.

## The Cancer Genome Atlas (TCGA)

TCGA is a comprehensive program in cancer genomics that is jointly supported and managed by the National Cancer Institute and the National Human Genome Research Institute of the U.S. National Institutes of Health. TCGA began in 2006 as a pilot focused on three projects, glioblastoma multiforme (GBM), serous cystadenocarcinoma of the ovary, and lung squamous carcinoma, and has recently expanded to produce comprehensive genomic data sets for at least 10 additional cancers in the next two years. Given TCGA's contributions in launching the ICGC and cooperation to ensure that its policies (posted at <http://cancergenome.nih.gov>) are coordinated with those of the ICGC, TCGA's participation in the ICGC is considered to be equivalent to that of a full member. TCGA, however, is not able to join the ICGC formally at this time, because of technical and legal issues in the U.S. related to the mechanisms of the distribution of controlled-access data, although such data are directly available to investigators at <http://cancergenome.nih.gov/dataportal>. The National Institutes of Health policies relating to distribution of controlled-access datasets are being reviewed with the intent of enabling researchers to integrate and analyze across databases, for example, using the franchise model adopted by the ICGC. Meanwhile, TCGA is ensuring that projects are coordinated and data sets are compatible with those of the Consortium.

## ICGC in the Next Decade

A large proportion of common cancers affecting patients around the world have been or will soon be selected for comprehensive cancer genome studies. Further efforts will be needed to leverage support and expertise to tackle the remaining tumor types, including rare cancers. The challenges of the ICGC are daunting due to the scope of the initiative, the complexity that is inherent to the heterogeneity of cancer and the limitations of current technologies to provide accurate long-range assemblies of highly rearranged chromosomes found in tumor



cells. These challenges underscore the importance of continued international coordination and further engagement of the scientific community in the next decade.

## Moving Towards Clinical Applications

ICGC catalogues, which are expected to grow exponentially, will have immediate relevance in the cancer research community. Early insight into the biology of somatic mutations will come from functional studies in cell-based and animal models of tumors. Mutation screens in retrospective tumor banks linked to registries or clinical trials having significant clinical data will inform on the potential clinical utility of somatic mutations as biomarkers for prognosis or drug-response. Germline variants identified by ICGC projects may allow the discovery of genes predisposing to familial malignancies, such as PALB2 and pancreatic cancer<sup>12,34</sup>. High throughput screens of RNAi or small molecule libraries, and the adaptation of existing model systems, will play a major role in refining potential therapeutic candidates for further study<sup>35</sup>.

Translating these discoveries into clinical practice will require more sophisticated clinical trials that take into account the increases in phenotypic subdivisions, additional coordination to identify subjects having tumors with similar profiles, and increased use of biomarkers, genomic analyses, informatics and other technologies in the clinical development of new therapeutics. Given the tremendous potential for relatively low-cost genomic sequencing to reveal clinically useful information, we anticipate that in the not so distant future, partial or full cancer genomes will routinely be sequenced as part of the clinical evaluation of cancer patients and as part of their on-going clinical management. The successful and appropriate translation of cancer genome research into clinical practice will raise important social and ethical questions. It will be essential to combine the expertise of oncologists, biostatisticians, pathologists, geneticists, policy-makers and members of the biopharmaceutical industry to meet this challenge by developing new policies and clinical paradigms that enable rapid translation of many new biomarkers and cancer targets into new clinical tests and therapeutic interventions that will benefit cancer patients.

### Box 1

#### Goals of the ICGC

- Coordinate the generation of comprehensive catalogues of genomic abnormalities (somatic mutations) in tumors in 50 different cancer types and/or subtypes which are of clinical and societal importance across the globe.
- Ensure high quality by defining the catalogue for each tumor type or subtype to include the full range of somatic mutations such as single-nucleotide variants, insertions, deletions, copy number changes, translocations and other chromosomal rearrangements, and to have the following features:
  - Comprehensiveness, such that most cancer genes with somatic abnormalities occurring at a frequency of greater than 3% are discovered;
  - High resolution, ideally at a single nucleotide level;



- High quality, using common standards for pathology and technology;
  - Data from matched non-tumor tissue, to distinguish somatic from inherited sequence variants and aberrations;
  - Generate complementary catalogues of transcriptomic and epigenomic datasets from the same tumors.
- Make the data available to the entire research community as rapidly as possible, and with minimal restrictions, to accelerate research into the causes and control of cancer.
  - Coordinate research efforts so that the interests and priorities of individual participants, self-organizing consortia, funding agencies and nations are addressed, including use of the burden of disease and the minimization of unnecessary redundancy in tumor analysis efforts.
  - Support the dissemination of knowledge and standards related to new technologies, software, and methods to facilitate data integration and sharing with cancer researchers around the globe.

## Box 2

### Core Bioethical Elements

For *prospective research*, ICGC members should convey to potential participants, that:

- The ICGC is a coordinated effort among related scientific research projects being carried on around the world
- Participation in the ICGC and its component projects is voluntary
- Samples and data collected will be used for cancer research, which may include whole genome sequencing
- The patient's care will not be affected by their decision regarding participation
- The samples collected will be in limited quantities; access to them will be tightly controlled and will depend on the policy and practices of the ICGC-member project. At least a small percentage of the samples may be shared with laboratories in other countries for the purposes of performing quality control studies
- Data derived from the samples collected and data generated by the ICGC members will be made accessible to ICGC members and other international researchers through either an open or a controlled access database under terms and conditions that will maximize participant confidentiality
- The researchers accessing data and samples will be required to affirm that they will not attempt to re-identify participants
- There is a remote risk of being identified from data available on the databases

- Once data are placed in open databases, those data cannot be withdrawn later
- In controlled access databases the links to (local) data that can identify an individual will be destroyed upon withdrawal. Data previously distributed will continue to be used
- ICGC members agree not to make claims to possible IP on primary data
- No profit from eventual commercial products will be returned to subjects donating samples

For *retrospective research*, the above guidelines remain the same, with the exception that where the individual is no longer a patient, there will not be a concern that their care could be affected by participation.

For research involving *samples and data from deceased individuals*:

- Where required by law or ethics, consent should always be obtained from the families of a deceased individual if their samples and data are to be used; if re-consent is not required, however, ethics review is sufficient
- Ethics committee review should be sought for all research proposing the use of existing sample and data collections
- Existing collections are a limited and valuable resource; access to them will be tightly controlled.

For research using *anonymized samples*, ethics review may be required in some jurisdictions.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank research participants who are generously donating samples and data, as well as physicians and clinical staff contributing to sample annotation and collection. A complete list of organizations that support ICGC projects accompanies this article online.

## ICGC Marker Paper Author List As of 10 March 2010

**International Cancer Genome Consortium** (Project leaders and committee chairs are listed first and are designated with an asterix; other principal investigators and collaborators are listed in alphabetical order or as designated by the project leader).

### Executive Committee

Thomas J. Hudson<sup>\*1,2</sup>, Warwick Anderson<sup>3</sup>, Axel Aretz<sup>4</sup>, Anna D. Barker<sup>5</sup>, Cindy Bell<sup>6</sup>, Rosa R. Bernabé<sup>7</sup>, M. K. Bhan<sup>8</sup>, Fabien Calvo<sup>9</sup>, Iiro Eerola<sup>10</sup>, Daniela S. Gerhard<sup>5</sup>, Alan Gutmacher<sup>11</sup>, Mark Guyer<sup>12</sup>, Fiona M. Hemsley<sup>13</sup>, Jennifer L. Jennings<sup>1</sup>, David Kerr<sup>14,15</sup>, Peter Klatt<sup>7</sup>, Patrik Kolar<sup>10</sup>, Jun Kusuda<sup>16</sup>, David P. Lane<sup>13</sup>, Frank Laplace<sup>17</sup>, Youyong Lu<sup>18</sup>, Gerd Nettekoven<sup>19</sup>, Brad Ozenberger<sup>12</sup>, Jane Peterson<sup>12</sup>, T.S. Rao<sup>8</sup>, Jacques Remacle<sup>10</sup>, Alan J. Schafer<sup>20</sup>, Tatsuhiro Shibata<sup>21</sup>, Michael R. Stratton<sup>22</sup>, Joseph G. Vockley<sup>5</sup>, Koichi Watanabe<sup>23</sup>, Huanming Yang<sup>24</sup>, Matthew M. F. Yuen<sup>25</sup>.

### **Ethics and Policy Committee**

Bartha M. Knoppers<sup>\*26</sup>, Martin Bobrow<sup>27</sup>, Anne Cambon-Thomsen<sup>28</sup>, Lynn G. Dressler<sup>29</sup>, Stephanie O. M. Dyke<sup>22</sup>, Yann Joly<sup>26</sup>, Kazuto Kato<sup>30</sup>, Karen L. Kennedy<sup>22</sup>, Pilar Nicolás<sup>31</sup>, Michael J. Parker<sup>32</sup>, Emmanuelle Rial-Sebbag<sup>28</sup>, Carlos M. Romeo-Casabona<sup>31</sup>, Kenna M. Shaw<sup>5</sup>, Susan Wallace<sup>26</sup>, Georgia L. Wiesner<sup>33,34</sup>, Nikolajs Zeps<sup>35,36</sup>.

### **Tissue and Clinical Annotation Working Group**

<sup>1</sup>Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada.

<sup>2</sup>Departments of Medical Biophysics and Molecular Genetics, University of Toronto, Toronto, ON M5S 1A1, Canada.

<sup>3</sup>National Health and Medical Research Council, Canberra, Australian Capital Territory 2601, Australia.

<sup>4</sup>Project Management Agency, German Aerospace Center (DLR), 53175 Bonn, Germany.

<sup>5</sup>National Cancer Institute, US National Institutes of Health, Bethesda, MD 20892, USA.

<sup>6</sup>Genome Canada, Ottawa, ON K2P 1P1, Canada.

<sup>7</sup>Secretariat of State for Research, Ministry of Science and Innovation, 28027 Madrid, Spain.

<sup>8</sup>Department of Biotechnology, Ministry of Science & Technology, Government of India, New Delhi, Delhi 110003, India.

<sup>9</sup>Institut National du Cancer, 92513 Boulogne-Billancourt, France.

<sup>10</sup>Genomics and Systems Biology Unit, Health Research Directorate, European Commission, B-1049 Brussels, Belgium.

<sup>11</sup>Eunice Kennedy Shriver National Institute of Child Health and Human Development, US National Institutes of Health, Bethesda, MD 20892, USA.

<sup>12</sup>National Human Genome Research Institute, US National Institutes of Health, Bethesda, MD 20892, USA.

<sup>13</sup>Cancer Research UK, London WC2A 3PX, UK.

<sup>14</sup>Sidra Medical and Research Center, Qatar Foundation, Doha, Qatar.

<sup>15</sup>Department of Clinical Pharmacology, University of Oxford, Oxford OX2 6HE, UK.

<sup>16</sup>National Institute of Biomedical Innovation, Ibaraki, Osaka 567-0085, Japan.

<sup>17</sup>Division of Molecular Life Sciences, Federal Ministry of Education and Research, 11055 Berlin, Germany.

<sup>18</sup>Beijing Cancer Institute and Hospital, Peking University School of Oncology, 100142 Beijing, China.

<sup>19</sup>German Cancer Aid, 53113 Bonn, Germany.

<sup>20</sup>Wellcome Trust, London NW1 2BE, UK.

<sup>21</sup>National Cancer Center Research Institute, Chuo-ku, Tokyo 104-0045, Japan.

<sup>22</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.

<sup>23</sup>Yokohama Institute, RIKEN, Yokohama, Kanagawa 230-0045, Japan.

<sup>24</sup>BGI-Shenzhen, Shenzhen, 518083 Guangdong, China.

<sup>25</sup>The Hong Kong University of Science and Technology, Hong Kong, China.

<sup>26</sup>Centre of Genomics and Policy, McGill University and Génome Québec Innovation Centre, Montreal, QC H3A 1A4, Canada.

<sup>27</sup>Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, UK.

<sup>28</sup>U558, INSERM, 31073 Toulouse, France.

<sup>29</sup>University of North Carolina School of Pharmacy, Division of Pharmaceutical Outcomes and Policy, Institute for Pharmacogenomics and Individualized Therapy, Chapel Hill, NC 27599, USA.

<sup>30</sup>Institute for Research in Humanities, Graduate School of Biostudies, Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Kyoto 606-8501, Japan.

<sup>31</sup>Inter-University Chair in Law and the Human Genome, University of Deusto, Bilbao, 48007 Bizkaia, Spain.

<sup>32</sup>The Ethox Centre, University of Oxford, Oxford OX3 7LF, UK.

<sup>33</sup>Department of Genetics, Case Western Reserve University, Cleveland, OH 44106, USA.

<sup>34</sup>Center for Human Genetics, University Hospitals Case Medical Center, Cleveland, OH 44106, USA.

<sup>35</sup>St John of God Pathology, Subiaco, Western Australia 6008, Australia.

<sup>36</sup>Schools of Surgery and Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, Western Australia 6009, Australia.

Peter Lichter<sup>\*37</sup>, Andrew V. Biankin<sup>38,39</sup>, Christian Chabannon<sup>9,40</sup>, Lynda Chin<sup>41,42</sup>, Bruno Clément<sup>43</sup>, Enrique de Alava<sup>44</sup>, Françoise Degos<sup>45</sup>, Martin L. Ferguson<sup>46</sup>, Peter Geary<sup>47</sup>, D. Neil Hayes<sup>48</sup>, Thomas J. Hudson<sup>1,2</sup>, Amber L. Johns<sup>38</sup>, Arek Kasprzyk<sup>1</sup>, Hidewaki Nakagawa<sup>49</sup>, Robert Penny<sup>50</sup>, Miguel A. Piris<sup>51</sup>, Rajiv Sarin<sup>52</sup>, Aldo Scarpa<sup>53,54</sup>, Tatsuhiro Shibata<sup>21</sup>, Marc van de Vijver<sup>55,56</sup>.

### Technologies Working Group

Michael R. Stratton<sup>\*22</sup>, Hiroyuki Aburatani<sup>57</sup>, Mónica Bayés<sup>58,59</sup>, David D.L. Bowtell<sup>60,61</sup>, Peter J. Campbell<sup>22,62</sup>, Xavier Estivill<sup>58,59</sup>, P. Andrew Futreal<sup>22</sup>, Daniela S. Gerhard<sup>5</sup>, Sean M. Grimmond<sup>63</sup>, Ivo Gut<sup>64</sup>, Martin Hirst<sup>65</sup>, Carlos López-Otín<sup>66</sup>, Partha Majumder<sup>67</sup>, Marco Marra<sup>65</sup>, John D. McPherson<sup>1,68</sup>, Hidewaki Nakagawa<sup>49</sup>, Zemin Ning<sup>22</sup>, Xose S. Puente<sup>66</sup>, Yijun Ruan<sup>69</sup>, Tatsuhiro Shibata<sup>21</sup>, Hendrik G. Stunnenberg<sup>70</sup>, Harold Swerdlow<sup>22</sup>, Victor E. Velculescu<sup>71</sup>, Richard K. Wilson<sup>72,73</sup>, Hong H. Xue<sup>74,75</sup>, Liu Yang<sup>76</sup>.

<sup>37</sup>German Cancer Research Center, 69120 Heidelberg, Germany.

<sup>38</sup>Garvan Institute of Medical Research, University of New South Wales, Darlinghurst, Sydney, New South Wales 2010, Australia.

<sup>39</sup>Department of Surgery, Bankstown Hospital, Bankstown, Sydney, New South Wales 2200, Australia.

<sup>40</sup>Institut Paoli-Calmettes, 13273 Marseille, France.

<sup>41</sup>Belfer Institute for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, MA 02115, USA.

<sup>42</sup>Department of Dermatology, Harvard Medical School, Boston, MA 02115, USA.

<sup>43</sup>U991, INSERM, 35043 Rennes, France.

<sup>44</sup>Department of Hematology, Centro de Investigación del Cáncer, Hospital Universitario, Universidad de Salamanca, 37007 Salamanca, Spain.

<sup>45</sup>Hôpital Beaujon, 92110 Clichy, France.

<sup>46</sup>MLF Consulting, Arlington, MA 02474, USA.

<sup>47</sup>Canadian Tumour Repository Network, Winnipeg, MN R3M 0V5, Canada.

<sup>48</sup>Department of Internal Medicine, Division of Medical Oncology, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA.

<sup>49</sup>Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa 230-0045, Japan.

<sup>50</sup>International Genomics Consortium, Phoenix, AZ 85004, USA.

<sup>51</sup>Molecular Pathology Programme, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain.

<sup>52</sup>Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, Maharashtra 410210, India.

<sup>53</sup>Department of Pathology, University of Verona, 37134 Verona, Italy.

<sup>54</sup>Center for Applied Research on Cancer (ARC-NET), Verona University Hospital, 37134 Verona, Italy.

<sup>55</sup>Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands.

<sup>56</sup>Academic Medical Center, 1015 AZ Amsterdam, The Netherlands.

<sup>57</sup>Research Center for Advanced Science and Technology, University of Tokyo, Meguro-ku, Tokyo 153-8904, Japan.

<sup>58</sup>Center for Genomic Regulation, Pompeu Fabra University, 08003 Barcelona, Spain.

<sup>59</sup>Public Health and Epidemiology Network Biomedical Research Center (CIBERESP), Barcelona, 08003 Catalonia, Spain.

<sup>60</sup>Peter MacCallum Cancer Centre, Melbourne, Victoria 3002, Australia.

<sup>61</sup>Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3010, Australia.

<sup>62</sup>Department of Haematology, University of Cambridge, Cambridge CB2 2XY, UK.

<sup>63</sup>Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4067, Australia.

<sup>64</sup>CEA/DSV/IG-Centre National de Genotypage, 91057 Evry, France.

<sup>65</sup>Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 1L3, Canada.

<sup>66</sup>Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología, Universidad de Oviedo, 33006 Oviedo, Spain.

<sup>67</sup>National Institute of Biomedical Genomics, Kalyani, West Bengal 741251, India.

<sup>68</sup>Department of Medical Biophysics, University of Toronto, Toronto, ON M5S 1A1, Canada.

<sup>69</sup>Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138672, Singapore.

<sup>70</sup>Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen, 6500 HB Nijmegen, The Netherlands.

<sup>71</sup>Ludwig Center for Cancer Genetics and Therapeutics, Johns Hopkins Kimmel Cancer Center, Baltimore, MD 21231, USA.

<sup>72</sup>The Genome Center, Washington University School of Medicine, St. Louis, MO 63108, USA.

<sup>73</sup>Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO 63108, USA.

<sup>74</sup>Applied Genomics Center, Fok Ying Tung Graduate School, HKUST, Hong Kong, China.

<sup>75</sup>Department of Biochemistry, The Hong Kong University of Science and Technology, Hong Kong, China.

<sup>76</sup>Cancer Institute, Zhejiang University, 310009 Hangzhou, China.

### Bioinformatics Analyses Working Group

Paul T. Spellman<sup>\*77</sup>, Gary D. Bader<sup>78,79</sup>, Paul C. Boutros<sup>1</sup>, Peter J. Campbell<sup>22,62</sup>, Paul Flicek<sup>80</sup>, Gad Getz<sup>81</sup>, Roderic Guigó<sup>82</sup>, Guangwu Guo<sup>24</sup>, David Haussler<sup>83</sup>, Simon Heath<sup>64</sup>, Tim J. Hubbard<sup>22</sup>, Tao Jiang<sup>24</sup>, Steven M. Jones<sup>65</sup>, Qibin Li<sup>24</sup>, Nuria López-Bigas<sup>84</sup>, Ruibang Luo<sup>24</sup>, Lakshmi Muthuswamy<sup>1</sup>, B. F. Francis Ouellette<sup>1</sup>, John V. Pearson<sup>63</sup>, Xose S. Puente<sup>66</sup>, Victor Quesada<sup>66</sup>, Benjamin J. Raphael<sup>85</sup>, Chris Sander<sup>86</sup>, Tatsuhiro Shibata<sup>21</sup>, Terence P. Speed<sup>87,88</sup>, Lincoln D. Stein<sup>1</sup>, Joshua M. Stuart<sup>89</sup>, Jon W. Teague<sup>22</sup>, Yasushi Totoki<sup>21</sup>, Tatsuhiko Tsunoda<sup>49</sup>, Alfonso Valencia<sup>90</sup>, David A. Wheeler<sup>91</sup>, Honglong Wu<sup>24</sup>, Shancen Zhao<sup>24</sup>, Guangyu Zhou<sup>24</sup>.

### Data Coordination and Management Working Group

Lincoln D. Stein<sup>\*1</sup>, Roderic Guigó<sup>82</sup>, Tim J. Hubbard<sup>22</sup>, Yann Joly<sup>26</sup>, Steven M. Jones<sup>65</sup>, Arek Kasprzyk<sup>1</sup>, Mark Lathrop<sup>64,92</sup>, Nuria López-Bigas<sup>84</sup>, B. F. Francis Ouellette<sup>1</sup>, Paul T. Spellman<sup>77</sup>, Jon W. Teague<sup>22</sup>, Gilles Thomas<sup>93,94</sup>, Alfonso Valencia<sup>90</sup>, Teruhiko Yoshida<sup>21</sup>.

### Data Release, Data Tiers and Publications Working Group

Karen L. Kennedy<sup>\*22</sup>, Myles Axton<sup>95</sup>, Stephanie O. M. Dyke<sup>22</sup>, P. Andrew Futreal<sup>22</sup>, Daniela S. Gerhard<sup>5</sup>, Chris Gunter<sup>96</sup>, Mark Guyer<sup>12</sup>, Thomas J. Hudson<sup>1,2</sup>, John D. McPherson<sup>1,68</sup>, Linda J. Miller<sup>97</sup>, Brad Ozenberger<sup>12</sup>, Kenna M. Shaw<sup>5</sup>.

### Data Coordination Centre

<sup>77</sup>Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94510, USA.

<sup>78</sup>Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada.

<sup>79</sup>Banting and Best Department of Medical Research, University of Toronto, Toronto, ON M5S 3E1, Canada.

<sup>80</sup>European Molecular Biology Laboratory-European Bioinformatics Institute, Hinxton, Cambridge CB10 1SD, UK.

<sup>81</sup>Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA.

<sup>82</sup>Spanish National Bioinformatics Institute (INB) and Center for Genomic Regulation, Universitat Pompeu Fabra, 08003 Barcelona, Spain.

<sup>83</sup>Howard Hughes Medical Institute and Center for Biomolecular Science and Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA.

<sup>84</sup>Research Unit on Biomedical Informatics, Department of Experimental and Health Science, Pompeu Fabra University, 08003 Barcelona, Spain.

<sup>85</sup>Department of Computer Science & Center for Computational Molecular Biology, Brown University, Providence, RI 02912, USA.

<sup>86</sup>Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.

<sup>87</sup>Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia.

<sup>88</sup>Department of Statistics, University of California Berkeley, Berkeley, CA 94720, USA.

<sup>89</sup>Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA.

<sup>90</sup>Spanish National Bioinformatics Institute (INB) and Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain.

<sup>91</sup>Human Genome Sequencing Center & Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA.

<sup>92</sup>Fondation Jean Dausset, Centre d'Etude du Polymorphisme Humain, 75010 Paris, France.

<sup>93</sup>Université Claude Bernard Lyon 1, 69622 Villeurbanne, France.

<sup>94</sup>Fondation Synergie Lyon Cancer, 69008 Lyon, France.

<sup>95</sup>Nature Genetics, New York, NY 10013-1917, USA.

<sup>96</sup>HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA.

<sup>97</sup>Nature and the Nature research journals, New York, NY 10013, USA.

Arek Kasprzyk<sup>\*1</sup>, Lincoln D. Stein<sup>\*1</sup>, Junjun Zhang<sup>1</sup>, Syed A. Haider<sup>98</sup>, Jianxin Wang<sup>1</sup>, Christina K. Yung<sup>1</sup>, Anthony Cross<sup>1</sup>, Yong Liang<sup>1</sup>, Saravanamuttu Gnaneshan<sup>1</sup>, Jonathan Guberman<sup>1</sup>, Jack Hsu<sup>1</sup>.

#### International Data Access Committee

Martin Bobrow<sup>\*27</sup>, Don R. C. Chalmers<sup>99</sup>, Karl W. Hasel<sup>6</sup>, Yann Joly<sup>26</sup>, Terry S. H. Kaan<sup>100</sup>, Karen L. Kennedy<sup>22</sup>, Bartha M. Knoppers<sup>26</sup>, William W. Lowrance<sup>101</sup>, Tohru Masui<sup>16</sup>, Pilar Nicolás<sup>31</sup>, Emmanuelle Rial-Sebbag<sup>28</sup>, Laura Lyman Rodriguez<sup>12</sup>, Catherine Vergely<sup>102</sup>, Teruhiko Yoshida<sup>21</sup>.

#### Australia – Pancreatic Cancer (Ductal adenocarcinoma) and Ovarian Cancer (Serous adenocarcinoma)

Sean M. Grimmond<sup>\*63</sup>, Andrew V. Biankin<sup>38,39</sup>, David D. L. Bowtell<sup>60,61</sup>, Nicole Cloonan<sup>63</sup>, Anna deFazio<sup>103,104</sup>, James R. Eshleman<sup>105</sup>, Dariush Etemadmoghadam<sup>60,61</sup>, Brooke A. Gardiner<sup>63</sup>, James G. Kench<sup>38,106</sup>, Aldo Scarpa<sup>53,54</sup>, Robert L. Sutherland<sup>38</sup>, Margaret A. Tempero<sup>107</sup>, Nicola J. Waddell<sup>63</sup>, Peter J. Wilson<sup>63</sup>.

#### Canada – Pancreatic Cancer (Ductal adenocarcinoma)

John D. McPherson<sup>\*1,68</sup>, Steve Gallinger<sup>108,109</sup>, Ming-Sound Tsao<sup>110,111</sup>, Patricia A. Shaw<sup>112</sup>, Gloria M. Petersen<sup>113</sup>, Debabrata Mukhopadhyay<sup>114</sup>, Lynda Chin<sup>41,42</sup>, Ronald A. DePinho<sup>41,115</sup>, Sarah Thayer<sup>116</sup>, Lakshmi Muthuswamy<sup>1</sup>, Kamran Shazand<sup>1</sup>, Timothy Beck<sup>1</sup>, Michelle Sam<sup>1</sup>, Lee Timms<sup>1</sup>, Vanessa Ballin<sup>1</sup>.

#### China – Gastric Cancer (Intestinal- and diffuse-type)

Youyong Lu<sup>\*18</sup>, Jiafu Ji<sup>18</sup>, Xiuqing Zhang<sup>24</sup>, Feng Chen<sup>18</sup>, Xueda Hu<sup>24</sup>, Guangyu Zhou<sup>24</sup>, Qi Yang<sup>24</sup>, Geng Tian<sup>24</sup>, Lianhai Zhang<sup>18</sup>, Xiaofang Xing<sup>18</sup>, Xianghong Li<sup>18</sup>, Zhenggang Zhu<sup>117</sup>, Yingyan Yu<sup>117</sup>, Jun Yu<sup>118</sup>, Huanming Yang<sup>24</sup>.

<sup>98</sup>Computer Laboratory, University of Cambridge, Cambridge CB3 0FD, UK.

<sup>99</sup>Faculty of Law, University of Tasmania, Hobart, Tasmania 7001, Australia.

<sup>100</sup>Faculty of Law, National University of Singapore, Singapore 259776, Singapore.

<sup>101</sup>Consultant in Health Research Ethics and Policy, 34280 La Grande Motte, France.

<sup>102</sup>ISIS 39 rue Camille Desmoulins, Institut Gustav Roussy, Pediatric Sce, 94805 Villejuif, France.

<sup>103</sup>Department of Gynaecological Oncology, Westmead Hospital, Westmead, Sydney, New South Wales 2145, Australia.

<sup>104</sup>Westmead Institute for Cancer Research, University of Sydney at the Westmead Millennium Institute, Westmead, Sydney, New South Wales 2145, Australia.

<sup>105</sup>Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore, MD 21231, USA.

<sup>106</sup>Department of Anatomical Pathology, Royal Prince Alfred Hospital, University of Sydney, Camperdown, Sydney, New South Wales 2050, Australia.

<sup>107</sup>Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94115, USA.

<sup>108</sup>Department of General Surgery, Toronto General Hospital, Toronto, ON M5G 2C4, Canada.

<sup>109</sup>Samuel Lunenfeld Research Institute, Toronto, ON M5S 1A1, Canada.

<sup>110</sup>Ontario Cancer Institute, University Health Network, Toronto, ON M5G 2M9, Canada.

<sup>111</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON M5S 1A1, Canada.

<sup>112</sup>Department of Pathology, University Health Network, Toronto, ON M5G 2C4, Canada.

<sup>113</sup>Department of Health Science Research, Mayo Clinic, Rochester, MN 55905, USA.

<sup>114</sup>Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55905, USA.

<sup>115</sup>Department of Medicine and Genetics, Harvard Medical School, Boston, MA 02115, USA.

<sup>116</sup>Department of Surgery, Harvard Medical School, Boston, MA 02115, USA.

<sup>117</sup>Shanghai Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

<sup>118</sup>Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China.



### **European Union/France – Renal Cancer (Renal cell carcinoma; focus on but not limited to clear cell subtype)**

Mark Lathrop<sup>\*64,92</sup>, Jörg Tost<sup>64,92</sup>, Paul Brennan<sup>119</sup>, Ivana Holcatova<sup>120</sup>, David Zaridze<sup>121</sup>, Alvis Brazma<sup>80</sup>, Lars Egevad<sup>122</sup>, Egor Prokhortchouk<sup>123</sup>, Rosamonde Elizabeth Banks<sup>124</sup>, Mathias Uhlén<sup>125</sup>, Anne Cambon-Thomsen<sup>28</sup>, Juris Viksna<sup>126</sup>, Fredrik Ponten<sup>127</sup>, Konstantin Skryabin<sup>128</sup>.

### **European Union/United Kingdom – Breast Cancer (Subtypes defined by an amplification of ER+ HER ductal-type)**

Michael R. Stratton<sup>\*22</sup>, P. Andrew Futreal<sup>22</sup>, Ewan Birney<sup>80</sup>, Ake Borg<sup>129</sup>, Anne-Lise Børresen-Dale<sup>130,131</sup>, Carlos Caldas<sup>132</sup>, John A. Foekens<sup>133</sup>, Sancha Martin<sup>22</sup>, Jorge S. Reis-Filho<sup>134</sup>, Andrea L. Richardson<sup>135,136</sup>, Christos Sotiriou<sup>137</sup>, Hendrik G. Stunnenberg<sup>70</sup>, Gilles Thomas<sup>93,94</sup>, Marc van de Vijver<sup>55,56</sup>, Laura van't Veer<sup>55</sup>.

### **France – Breast Cancer (Subtype defined by an amplification of the HER2 gene)**

Fabien Calvo<sup>\*9</sup>, Daniel Birnbaum<sup>40</sup>, Hélène Blanche<sup>92</sup>, Pascal Boucher<sup>9</sup>, Sandrine Boyault<sup>138</sup>, Christian Chabannon<sup>9,40</sup>, Ivo Gut<sup>64</sup>, Jocelyne D. Masson-Jacquemier<sup>40</sup>, Mark Lathrop<sup>64,92</sup>, Iris Pauporté<sup>9</sup>, Xavier Pivot<sup>139</sup>, Anne Vincent-Salomon<sup>140</sup>, Eric Tabone<sup>138</sup>, Charles Theillet<sup>141</sup>, Gilles Thomas<sup>93,94</sup>, Jörg Tost<sup>64,92</sup>, Isabelle Treilleux<sup>138</sup>.

### **France – Liver Cancer (Hepatocellular carcinoma; secondary to alcohol and adiposity)**

<sup>119</sup>International Agency for Research on Cancer, 69372 Lyon, France.

<sup>120</sup>Institute of Hygiene and Epidemiology, First Faculty of Medicine, Charles University in Prague, 121 08 Prague, Czech Republic.

<sup>121</sup>Department of Epidemiology and Prevention, N. N. Blokhin Russian Cancer Research Centre, Moscow 115478, Russian Federation.

<sup>122</sup>Karolinska Institutet, Karolinska University Hospital, SE-171 76 Stockholm, Sweden.

<sup>123</sup>Bioengineering Center, Russian Academy of Sciences, Moscow 117312, Russian Federation.

<sup>124</sup>Cancer Research UK Centre, Leeds Institute for Molecular Medicine, St James's University Hospital, Leeds LS9 7TF, UK.

<sup>125</sup>Science for Life Laboratory, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

<sup>126</sup>Institute of Mathematics and Computer Science, University of Latvia, Riga LV-1459, Latvia.

<sup>127</sup>Uppsala University, SE-751 05 Uppsala, Sweden.

<sup>128</sup>Kurchatov Scientific Center, Moscow 123182, Russian Federation.

<sup>129</sup>Department of Oncology, Lund University, SE-221 85 Lund, Sweden.

<sup>130</sup>Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, 0310 Oslo, Norway.

<sup>131</sup>Faculty of Medicine, University of Oslo, 0316 Oslo, Norway.

<sup>132</sup>Department Oncology, University of Cambridge and Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge CB2 0RE, UK.

<sup>133</sup>Department of Medical Oncology, Erasmus MC Rotterdam, Josephine Nefkens Institute and Cancer Genomics Centre, 3015 CE Rotterdam, The Netherlands.

<sup>134</sup>Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London SW3 6JB, UK.

<sup>135</sup>Dana-Farber Cancer Institute, Boston, MA 02115, USA.

<sup>136</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA.

<sup>137</sup>Jules Bordet Institute, B-1000 Brussels, Belgium.

<sup>138</sup>Centre Léon Bérard, 69373 Lyon, France.

<sup>139</sup>Hôpital Jean Minjoz, 25030 Besançon, France.

<sup>140</sup>Institut Curie, 75231 Paris, France.

<sup>141</sup>Centre Val d'Aurelle Paul-Lamarque, 34298 Montpellier, France.



Fabien Calvo<sup>\*9</sup>, Paulette Bioulac-Sage<sup>142</sup>, Bruno Clément<sup>43</sup>, Thomas Decaens<sup>143,144</sup>, Françoise Degos<sup>45</sup>, Dominique Franco<sup>145</sup>, Ivo Gut<sup>64</sup>, Marta Gut<sup>92</sup>, Simon Heath<sup>64</sup>, Mark Lathrop<sup>64,92</sup>, Didier Samuel<sup>146,147</sup>, Gilles Thomas<sup>93,94</sup>, Jessica Zucman-Rossi<sup>148</sup>.

#### **Germany – Pediatric Brain Tumors (Medulloblastoma, Pediatric Pilocytic Astrocytoma)**

Peter Lichter<sup>\*37</sup>, Roland Eils<sup>\*37,149</sup>, Benedikt Brors<sup>37</sup>, Jan O. Korbel<sup>80,150</sup>, Andrey Korshunov<sup>151</sup>, Pablo Landgraf<sup>152</sup>, Hans Lehrach<sup>153</sup>, Stefan Pfister<sup>37,154</sup>, Bernhard Radlwimmer<sup>37</sup>, Guido Reifenberger<sup>155</sup>, Michael D. Taylor<sup>156,157</sup>, Christof von Kalle<sup>158,159</sup>.

#### **India – Oral Cancer (Gingivobuccal)**

Partha P. Majumder<sup>\*67</sup>, Rajiv Sarin<sup>52</sup>, T. S. Rao<sup>8</sup>, M. K. Bhan<sup>8</sup>.

#### **Italy – Rare Pancreatic Tumors (Enteropancreatic endocrine tumors and rare pancreatic exocrine tumors; intraductal papillary mucinous neoplasms, solid pseudopapillary tumors, mucinous cystic neoplasms and other rarer tumors)**

Aldo Scarpa<sup>\*53,54</sup>, Paolo Pederzoli<sup>160</sup>, Rita T. Lawlor<sup>54</sup>, Massimo Delledonne<sup>161</sup>, Alberto Bardelli<sup>162,163</sup>, Andrew V. Biankin<sup>38,39</sup>, Sean M. Grimmond<sup>63</sup>, Thomas Gress<sup>164</sup>, David Klimstra<sup>165</sup>, Giuseppe Zamboni<sup>53</sup>.

#### **Japan – Liver Cancer (Hepatocellular carcinoma; virus associated)**

Tatsuhiro Shibata<sup>\*21</sup>, Yusuke Nakamura<sup>49,166</sup>, Hidewaki Nakagawa<sup>49</sup>, Jun Kusuda<sup>16</sup>, Tatsuhiko Tsunoda<sup>49</sup>, Satoru Miyano<sup>166</sup>, Hiroyuki Aburatani<sup>57</sup>, Kazuto Kato<sup>30</sup>, Akihiro Fujimoto<sup>49</sup>, Teruhiko Yoshida<sup>21</sup>.

<sup>142</sup>Hôpital Pellegrin, 33076 Bordeaux, France.

<sup>143</sup>Hôpital Henri Mondor, 94010 Créteil, France.

<sup>144</sup>U955, INSERM, 94000 Créteil, France.

<sup>145</sup>Hôpital Antoine Bécélère, 92141 Clamart, France.

<sup>146</sup>Centre Hepato-Bilaire, AP-HP Hôpital Paul-Brousse, 94800 Villejuif, France.

<sup>147</sup>U785, INSERM, 94800 Villejuif, France.

<sup>148</sup>U674, INSERM, 75010 Paris, France.

<sup>149</sup>BioQuant, Heidelberg University, 69120 Heidelberg, Germany.

<sup>150</sup>Genome Biology Unit, European Molecular Biology Laboratory, 69126 Heidelberg, Germany.

<sup>151</sup>Department of Neuropathology, Heidelberg University Hospital, 69120 Heidelberg, Germany.

<sup>152</sup>Clinic for Pediatric Oncology, Hematology and Immunology, Heinrich-Heine University Hospital, 40225 Düsseldorf, Germany.

<sup>153</sup>Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany.

<sup>154</sup>Department of Pediatric Hematology and Oncology, Heidelberg University Hospital, 69120 Heidelberg, Germany.

<sup>155</sup>Institute of Neuropathology, Heinrich-Heine University, 40001 Düsseldorf, Germany.

<sup>156</sup>Division of Neurosurgery, Hospital for Sick Children, Toronto, ON M5G 1X8, Canada.

<sup>157</sup>Labatt Brain Tumor Research Center, Hospital for Sick Children, Toronto, ON M5G 1X8, Canada.

<sup>158</sup>National Center for Tumor Diseases, 69120 Heidelberg, Germany.

<sup>159</sup>Division of Translational Oncology, German Cancer Research Center, 69120 Heidelberg, Germany.

<sup>160</sup>Department of Surgery, University Hospital Trust of Verona, 37134 Verona, Italy.

<sup>161</sup>Functional Genomics Center, Department of Biotechnology, University of Verona, 37134 Verona, Italy.

<sup>162</sup>Laboratory of Molecular Genetics, Institute for Cancer Research and Treatment, University of Torino, 10060 Torino, Italy.

<sup>163</sup>FIRC Institute of Molecular Oncology, 20139 Milan, Italy.

<sup>164</sup>Department of Gastroenterology, Endocrinology, Metabolism and Infectiology, University of Marburg, 35043 Marburg, Germany.

<sup>165</sup>Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.

<sup>166</sup>Human Genome Center, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan.

## Spain – Chronic Lymphocytic Leukemia (with mutated and unmutated IgVH)

Elias Campo<sup>\*167</sup>, Carlos López-Otín<sup>66</sup>, Xavier Estivill<sup>58,59</sup>, Roderic Guigó<sup>82</sup>, Silvia de Sanjosé<sup>168</sup>, Miguel A. Piris<sup>51</sup>, Emili Montserrat<sup>167</sup>, Marcos González-Díaz<sup>44</sup>, Xose S. Puente<sup>66</sup>, Pedro Jares<sup>167</sup>, Alfonso Valencia<sup>90</sup>, Heinz Himmelbaue<sup>58</sup>, Victor Quesada<sup>66</sup>, Silvia Bea<sup>167</sup>.

## United Kingdom – Breast Cancer (Triple Negative/lobular/other)

Michael R. Stratton<sup>22</sup>, P. Andrew Futreal<sup>22</sup>, Peter J. Campbell<sup>22,62</sup>, Anne Vincent-Salomon<sup>140</sup>, Andrea L. Richardson<sup>135,136</sup>, Jorge S. Reis-Filho<sup>134</sup>, Marc van de Vijver<sup>55,56</sup>, Gilles Thomas<sup>93,94</sup>, Jocelyne D. Masson-Jacquemier<sup>40</sup>, Samuel Aparicio<sup>169</sup>, Ake Borg<sup>129</sup>, Anne-Lise Børresen-Dale<sup>130,131</sup>, Carlos Caldas<sup>132</sup>, John A. Foekens<sup>133</sup>, Hendrik G. Stunnenberg<sup>70</sup>, Laura van't Veer<sup>55</sup>, Douglas F. Easton<sup>170</sup>, Paul T. Spellman<sup>77</sup>, Sancha Martin<sup>22</sup>.

## United States – The Cancer Genome Atlas

The Cancer Genome Atlas Research Network.

## Initial Scientific Planning Committee

Thomas J. Hudson<sup>\*1,2</sup>, Lynda Chin<sup>\*41,42</sup>, Bartha M. Knoppers<sup>\*26</sup>, Eric S. Lander<sup>\*81</sup>, Peter Lichter<sup>\*37</sup>, Lincoln D. Stein<sup>\*1</sup>, Michael R. Stratton<sup>\*22</sup>, Warwick Anderson<sup>3</sup>, Anna D. Barker<sup>5</sup>, Cindy Bell<sup>6</sup>, Martin Bobrow<sup>27</sup>, Wylie Burke<sup>171</sup>, Francis S. Collins<sup>172</sup>, Carolyn C. Compton<sup>5</sup>, Ronald A. DePinho<sup>41,115</sup>, Douglas F. Easton<sup>170</sup>, P. Andrew Futreal<sup>22</sup>, Daniela S. Gerhard<sup>5</sup>, Anthony R. Green<sup>173</sup>, Mark Guyer<sup>12</sup>, Stanley R. Hamilton<sup>174</sup>, Tim J. Hubbard<sup>22</sup>, Olli P. Kallioniemi<sup>175</sup>, Karen L. Kennedy<sup>22</sup>, Timothy J. Ley<sup>72,176</sup>, Edison T. Liu<sup>69</sup>, Youyong Lu<sup>18</sup>, Partha Majumder<sup>67</sup>, Marco Marra<sup>65</sup>, Brad Ozenberger<sup>12</sup>, Jane Peterson<sup>12</sup>, Alan J. Schafer<sup>20</sup>, Paul T. Spellman<sup>77</sup>, Hendrik G. Stunnenberg<sup>70</sup>, Brandon J. Wainwright<sup>177</sup>, Richard K. Wilson<sup>72,73</sup>, Huanming Yang<sup>24</sup>.

## References

1. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. 2009; 458:719–724. [PubMed: 19360079]
2. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000; 100:57–70. [PubMed: 10647931]

<sup>167</sup>Hospital Clínic, University of Barcelona, 08036 Barcelona, Spain.

<sup>168</sup>Unit of Infections and Cancer, Cancer Epidemiology Research Programme, CIBER Epidemiología y Salud Pública, Institut Català d'Oncologia-IDIBELL, 08907 Hospitalet de Llobregat, Spain.

<sup>169</sup>BC Cancer Research Centre, BC Cancer Agency, Vancouver, BC V5Z 1L3, Canada.

<sup>170</sup>Departments of Public Health and Primary Care and Oncology, University of Cambridge, Cambridge CB1 8RN, UK.

<sup>171</sup>Department of Bioethics and Humanities, University of Washington, Seattle, WA 98195, USA.

<sup>172</sup>US National Institutes of Health, Bethesda, MD 20892, USA.

<sup>173</sup>Cambridge Institute for Medical Research and Department of Haematology, University of Cambridge, Cambridge CB2 2XY, UK.

<sup>174</sup>Pathology and Laboratory Medicine, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA.

<sup>175</sup>Institute for Molecular Medicine Finland, University of Helsinki, FIN-00290 Helsinki, Finland.

<sup>176</sup>Departments of Medicine and Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA.

<sup>177</sup>Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia.

3. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. 2001; 344:783–792. [PubMed: 11248153]
4. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med*. 2001; 344:1031–1037. [PubMed: 11287972]
5. Druker BJ, Sawyers CL, Kantarjian H, Resta DJ, Reese SF, et al. Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N Engl J Med*. 2001; 344:1038–1042. [PubMed: 11287973]
6. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*. 2010; 463:184–190. [PubMed: 20016488]
7. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature*. 2010; 463:191–196. [PubMed: 20016485]
8. Greenman C, Stephens P, Smith R, Dalgleish GL, Hunter C, et al. Patterns of somatic mutation in human cancer genomes. *Nature*. 2007; 446:153–158. [PubMed: 17344846]
9. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. 2008; 455:1061–1068. [PubMed: 18772890]
10. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008; 455:1069–1075. [PubMed: 18948947]
11. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, et al. The genomic landscapes of human breast and colorectal cancers. *Science*. 2007; 318:1108–1113. [PubMed: 17932254]
12. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*. 2008; 321:1801–1806. [PubMed: 18772397]
13. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science*. 2008; 321:1807–12. [PubMed: 18772396]
14. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, et al. High frequency of mutations of the PIK3CA gene in human cancers. *Science*. 2004; 304:554. [PubMed: 15016963]
15. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, et al. Mutations of the BRAF gene in human cancer. *Nature*. 2002; 417:949–954. [PubMed: 12068308]
16. van Haafden G, Dalgleish GL, Davies H, Chen L, Bignell G, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat Genet*. 2009; 41:521–523. [PubMed: 19330029]
17. Dalgleish GL, Furge KL, Greenman C, Chen L, Bignell GR, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature*. 2010; 463:360–363. [PubMed: 20054297]
18. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, Kos I, Batinic-Haberle I, Jones S, Riggins GJ, Friedman H, Friedman A, Reardon D, Herndon J, Kinzler KW, Velculescu VE, Vogelstein B, Bigner DD. IDH1 and IDH2 mutations in gliomas. *N Engl J Med*. 2009; 360:765–773. [PubMed: 19228619]
19. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008; 456:66–72. [PubMed: 18987736]
20. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med*. 2009; 361:1058–1066. [PubMed: 19657110]
21. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature*. 2009; 461:809–813. [PubMed: 19812674]
22. Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature*. 2009; 462:1005–1010. [PubMed: 20033038]
23. Dickson D. Wellcome funds cancer database. *Nature*. 1999; 401:729. [PubMed: 10548085]

24. Collins FS, Barker AD. Mapping the cancer genome. Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies. *Sci Am.* 2007; 296:50–57. [PubMed: 17348159]
25. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008; 5:621–628. [PubMed: 18516045]
26. Shah SP, Köbel M, Senz J, Morin RD, Clarke BA, et al. Mutation of FOXL2 in granulosa-cell tumors of the ovary. *N Engl J Med.* 2009; 360:2719–2729. [PubMed: 19516027]
27. Haider S, Ballester B, Smedley D, Zhang J, Rice P, et al. BioMart Central Portal--unified access to biological data. *Nucleic Acids Res.* 2009; 37:W23–W27. [PubMed: 19420058]
28. Hubbard TJ, Aken BL, Ayling S, Ballester B, Beal K, et al. Ensembl 2009. *Nucleic Acids Res.* 2009; 37:D690–D697. [PubMed: 19033362]
29. Pruitt KD, Harrow J, Harte RA, Wallin C, Diekhans M, et al. The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res.* 2009; 19:1316–1323. [PubMed: 19498102]
30. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001; 29:308–311. [PubMed: 11125122]
31. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* 2007; 449:851–861. [PubMed: 17943122]
32. Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, et al. Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* 2009; 37:D619–622. [PubMed: 18981052]
33. Toronto International Data Release Workshop Authors. Prepublication data sharing. *Nature.* 2009; 461:168–170. [PubMed: 19741685]
34. Jones S, Hruban RH, Kamiyama M, Borges M, Zhang X, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science.* 2009; 324:217. [PubMed: 19264984]
35. Chin L, Gray JW. Translating insights from the cancer genome into clinical practice. *Nature.* 2008; 452:553–563. [PubMed: 18385729]